(12)

The Television as Robot Servant

Robert Thibadeau, Ph.D.

August, 1993
CMU-RI-TR-93-22

The Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

This research was sponsored in part by the Advanced Research Projects Agency (DOD) ARPA order 6873, under contract #MDA972-92-J-1010. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the US Government.

ABSTRACT

Television and computing have an interesting intersection in the topic of how to put computer software into the television infrastructure. It is widely recognized that the result is a home information appliance. However, despite previous commercial experimentation, such an appliance has not emerged as successful. Perhaps there is only a deficiency in properly conceptualizing the appliance. In keeping with this theme, this paper presents a theory of why people value information. Using this theory as a design aid in the creation of future home information appliances should improve their chances of commercial success.

The Television as Robot Servant

Robert Thibadeau, Ph.D.
School of Computer Science
Carnegie Mellon University

I walked into the family room and asked the TV if anything interesting happened today. It said that there is going to be a township meeting tonight that is going to discuss raising the school tax retroactively. Thought you might be interested.

Next morning, said "hi" to the TV. It said a semi-truck jackknifed on the expressway into work. A better route would be through Norcross.

That night, told the TV I need to eat out alone. It said the Italian place down the street was having a special on subs -- half price. No other recommendations.

If television sets get smart, isn't this the best way? Television attracts us because it services our need for entertainment, sports, and news. Could it not also service our need for other kinds of information? This question is old, and has been addressed by Sunday morning programming, PBS, C-SPAN and Court TV. But the question takes on new meaning when we ask if digital interactive (smart) TV can service our need for other kinds of information. Others have searched for the "killer" information service which will guarantee the success of smart TV. There may not be a single "killer" information service in the new age, but there may be principles that govern what works and what does not work. We think the right place to start is to answer the question "why is information valued?"

News conveys information that is valued because it is timely. Information also has value because it is pertinent to the locale and to personal needs. Television is a great servant because it pleasantly conveys pertinent information. The information that it currently conveys is a small subset of the information that people desire in day to day living. Our group has asked how far one can extend the information range of TV and remain entrancing. The answer is surprising and not quite as impossible as it sounds.

The current work was begun in late 1988 and early 1989 as an investigation of how to add "software to television." It had become obvious that television sets, within a decade, would become more computer-like in composition. For dozens of reasons, this was inevitable. For one, the delivery of multiple, compressed, digital video channels, a necessity in conserving precious electromagnetic spectrum space, meant that TVs were going to contain processors and memories. But using the computer to simply service the video programming requirements of television seems a grotesque under-utilization of the smart TV. We ask how to maximize that utilization without demanding too much of the rest of the infrastructure. The idea is to enhance the functionality of future computer software in the TV.

To a group such as ours that has been building systems capable of highly skilled perception, the answer, on its surface, is fairly obvious. Equip the television with knowledge of "who, what, and where it is," and then let it "watch" its own environment and "converse" with the humans out there. The television's environment would, naturally, be the cable feed (or satellite, broadcast, or video dial tone feed). "Conversing" could be through speech recognition and voice synthesis, and it could be through remote control buttons, text, and graphics. The television set then becomes the robot servant: like a Star Wars C3PO without arms or legs.

AVOIDING PAST MISTAKES The smart TV is a very different concept from the "computer." It is strictly an information servant. A computer is an information tool that requires skill to operate. It takes work to operate. Such an information tool on the television set seems quite wrong headed. In fact, such a thing has been tried many times before. The original information tool was videotex. This constantly re-broadcast some number of "pages" or screens. Usually only a few hundred screens are rebroadcast every few seconds. With enough effort, it is possible to find a screen one is looking for. Tragically, videotex classically contains very little, if any, information of value. Broadcasting the same thing to everybody, in little screens, means that, basically, nobody is getting much of any relevance to them. Regular video news is harder hitting, you can get some information just from the pictures, and you don't have to search for the information. It is presented to you.

Some people have decided the problem is with "text services," as if real people don't read. However, we suspect that this is not the problem. We suspect the problem is that the information is, first, not presented in a humanly simple fashion, and, second, nearly totally irrelevant.

Smart TV means tiered interactive services. Interactivity between a consumer and his TV will be local and telecommunicated. The video game is a supreme example of pure local interactivity that has gained huge acceptance. Another locally interactive medium, CD ROM, is on the rise. Locally interactive services do engage people in ways they are willing to buy.

Telecommunicated interactivity, like automated buying services, has not been widely accepted. Like videotex, perhaps this is due to requiring too much viewer involvement.

Being driven by immediate demands of business, the experts look to home shopping services and other means of getting people to impulse-buy in directions they can incrementally charge for. Furthermore, they look to places where money is already being spent by people, in their homes, on their TVs. They look to videotape sales and rentals with the resulting "pay-per-view" conception, and to video games with the resulting "Sega Channel" and "Nintendo Channel." The information highway owners want new things to put on their highways and new ways to charge tolls. This activity can easily put too many demands on the viewer.

The television industry has been loath to relinquish control to the audience. Very powerful home TV boxes are being created by Hewlett-Packard, Panasonic, Silicon Graphics, Phillips, General Electric, General Instruments, Scientific Atlanta, and IBM. Yet companies persist in wanting to control or, as bad, monitor every action of the consumer.

We believe the smart TV should provide service without the details of the service being monitored by the provider. In cable, the "addressable decoder box," promulgated by the cable box makers General Instruments and Scientific Atlanta, implements the vision of a TV as a remote terminal for a large scale mainframe. Of course, the "mainframe" is really a "headend" here, but the idea is the same. It does not follow that absolute control, all control except detailed show-product selecting, should reside in the headend.

Contrary to the popular view, giving up absolute control does not give up data security. Security does not require addressability. We believe smart TVs can be smart about what they can know and can be trusted to be responsible, like C3PO. Furthermore, there is no need to give up system control to software companies.

An interesting model of how smart TVsoftware might work is that of the Free Software Foundation in Boston. They produce great software that is distributed free and in source form. We do the same thing at CMU with the MACH operating system. FSF software is "copylefted," meaning that if you modify their code, they own it, and they are going to give your modifications away in source form. This naturally generates standards that are seriously open because the basic system and tool software is complete and free.

PROVIDING USEFUL INFORMATION The essence of the servant is that it assists without asking. Things are accomplished without thinking to accomplish them because the servant does the thinking. Thinking and making decisions are things people do not like to do. If you have to think, to make a decision, then what is the servant really for? Perhaps the whole crux of the matter is how to architect a smart TV system in such a way that decisions, thinking, is minimized while the value of the information is maximized. In descriptive algebra, we can define the simplest form of a utility function:

$$U = I / T$$

The Utility (or value) is the Information value divided by the amount of Thinking (number of decisions). Interestingly, I and T can be expressed in exactly the same units of information value (bits or, alternatively, "yes-no" decisions). In this case, utility is a dimensionless quantity because it divides bits of effective information value (the elimination of uncertainty) by the information value of the current decision (the current elimination of uncertainty).

This function can logically be rewritten purely in a time sense. Information uncertainty was reduced at some prior time that reduces the uncertainty now and therefore requires fewer decisions now. Put simply, we have:

$$U_{t_n} = I_{t_{<n}} / T_{t_n}$$

The utility depends strictly on what decisions are made for you prior to the time, t_n, that you make a decision. For convenience, returning to the first form of this equation, we can give Utility a unit value in bits of information value by simply squaring the numerator:

$U = I^2 / T$

This says, also, that the value of past decisions in utility plays an exponential role compared to the value of the present decision. The term need not be squared. It would be interesting to know:

$U = I^k / T$

where **k** is unknown and must be discovered by experimentation. This, more general, utility function would be preferred if we consider utility, U, to carry a reduction of uncertainty. Many people have noted that the value of the cellular telephone, TV news, pagers, and the like, is the reduction in uncertainty. Reduction in uncertainty is a gut thing, not just a mathematical one. One could argue, we believe, that most entertainment value is profitably viewed as reduction of uncertainty (creating and resolving it). Certainly, this is true of narrative structure, and narrative structure is generally regarded as the definitive form of storytelling. Music and other art forms have similar structures of creation and resolution.

In this model, people figure out how valuable information is to them -- always. There is always at least one decision (or a fraction of one) that pertains to the utility of the current information (since a divide by zero is not allowed.) Furthermore, utility can zoom to near infinity. It is possible to catch you exactly right. In other words, it is possible to create utility that is hugely valuable if not priceless.

We believe that the value of television is in a high value utility function: the simplicity of the decision in channel changing and then simply watching. It is well-known that most of the time most people randomly flick channels. I propose this means the system is creating fractional denominators (thinking decisions) that, by the formula, means high utility value even with relatively poor information. Another interesting experimental number might be the median density in conscious (non-motor) decisions per hour that people make during their waking hours. I suspect that number is low enough that many "experts" on "interactive TV" would have to think twice, and perhaps revise their revenue projections.

The summary of this speculative analysis is that utility in interactive TV comes from minimizing viewer decisions while maximizing information quality from prior decisions. The prior decisions can come from others or from the viewer. A prior decision made by oneself is the setup cost for the information service. An example is the decisions necessary to find and get an information service (e.g., Macy's Home Shopper) that one will later use.

If decisions come from others, they must be broadcast to be efficient and economical. In other words, one decision by one person must be the decision for more than one other person. This further modifies the equation (again, eliminating the time and exponential notation for simplicity in presentation):

$U = (I_{self} + I_{other}) / T$

The decisions by the other must replace decisions by the self to be counted here. The efficiency

of the system is enhanced if the I_{other} is itself the result of a broadcast multiplication factor, **B**, the bigger the better:

$$I'_{other} = I_{other} / B$$

or

$$U = (I_{self} + B\, I'_{other}) / T$$

In other words, one I'_{other} decision is multiplied by the number of people it effects. The larger the number, the more efficient the system is in reducing the "overall decision loading" and thus the service cost of interactive TV.

Several senior cable, telephone, and computer people have warned sternly against the use of the term "broadcast." However, our use is generic and should not be thought of as necessarily meaning airwaves. Indeed, magazines, newspapers, and CD ROM publishers are broadcasters -- as are all cable operators and the networks that inhabit the cables. We find the word "network" hard to swallow since "network" implies interaction (as in "networking"). The factor **B** represents the power of broadcasting, not of networking in general.

This is still only part of the story. There are other forms of decisions besides those that are laboriously made by others and those laboriously made by oneself. There is a derivative multiplier in the automaton. This is the idea of a robot actively making a decision not explicitly anticipated by the creator of the robot. Call this an automaton factor, **A.** This applies to the decisions that are involved in programming the automaton and those made by the automaton. The automaton factor is also a multiplier. Thus,

$$U = A\, (I_{self} + B\, I_{other}) / T$$

And, as we said earlier, the automaton factor will be exponential if we assume the utility, the value to the individual, is itself a reduction in uncertainty.

So, lets put a little meat on this analysis of how hard it is to create the smart TV. Our analysis says that people will get more utility out of situations that (a) minimize the current decision requirement, (b) utilize the decision requirements of others that are broadcast to them, (c) utilize as many past decisions of themselves (without adding to the historical decision load), and (d) utilize automatons that can leverage decisions. Furthermore, the argument is that this analysis is complete and deals with entertainment, news, or any useful "information event."

Actually, the analysis is not complete. There are other assumptions not revealed in this utility function. One has to do with the degree, as opposed to the tier level, of interactivity. Degree of interactivity, we believe, has to do with the responsiveness of a system: the time to wait for an expected response. Non-interactive systems in which a wait is required are less useful than those that provide near instant response. However, systems with predictable delays are not perceived

as bad. The "broadcaster's contract" is a contract to supply a program at a stated time and place. As long as people know when to expect something, and can reliably get it at the promised time and place, the utility is not greatly diminished. We assume, through tiered interactivity, that there will, in general, be either negligible waits or established broadcaster's contracts.

SOME PRACTICAL INSTANCES A low utility results from the use of a CD-ROM Encyclopedia for a high uncertainty search. This is where the person does not know exactly what he wants. For example, he knows there was a famous jurist in 1920 who made some clever remarks. He wants to check the remarks to find out if they would be appropriate for a talk he is giving on freedom of speech.

A much better utility for the CD-ROM Encyclopedia is in a low uncertainty search. The thinking (denominator) is much smaller. The person already knows the man's name and just checks the entry for "Oliver Wendell Holmes" for the saying "a clear and present danger."

We would argue that there are a great number of high uncertainty search problems in people's lives. These include simple questions like "Is the traffic bad from here to work this morning?" The uncertainty is not in how to ask the question, but who to ask. Who knows the answer? Another one is "What is murder like?," "What is sex like?," or "What is war like?" These are high uncertainty search problems because of difficulty in seeing how to gain the understanding. The entertainment industry has solved these search problems and really helped people explore the answers through character development and plot. Anybody can be a Greek god nowadays and watch those puny mortals on TV. Soon we may be able to control their puny fates! Here the high uncertainty problem has a low cost decision to watch and a great deal of information value in the revelations made by others and made affordable by broadcast.

The automaton, or what we have called, the friendly robot on the home TV, is a watcher whose environment is a high volume broadcast information stream. Pagers are well accepted watchers but they have only rudimentary intelligence and respond to lower volume broadcast information streams. The watcher works because a person can assign it a duty to watch out for something. For example, it could be set up to watch for problems in getting to work in the morning. Then, any time that it generates a useful event, it is generating a high information value event with only some set up (I_{self}), that was perhaps done over a year prior, and a decision (T) to watch a channel, or perhaps any channel, on TV. Or even easier, just to look and check out if the message light is on (so T=1 bit).

Another automaton type is a gopher that actively seeks information of value out of large databases. In the example of the encyclopedia, one simply gives the order "find the name and quips of the jurist in 1920 who was famous on the topic of freedom." The automaton itself has to develop appropriate search strategies on the encyclopedia. This type of gopher is not available in the foreseeable future. However, a gopher that automatically wakes up at 7 am and checks the traffic database for a possible problem going to work could perfectly emulate the watcher.

Gophers break down when there is a broadcast event. All gophers going for the same database at exactly the same time would not be good. So, for example, the winner of the football game or the

movement of a hurricane are best monitored by watchers, not gophers.

There are two sources of input in interactive TV. One is the data source (CD or cable), and the other is the viewer input (remote control). Intelligence can be applied to both.

The data source will be filtered and composed. In effect only some information from the source will be interesting. Data composition is not well studied but permits inference. The simplest form is to count instances of a message of interest as in "found 5 cars in the category you are looking for." More interesting is a composition that finds a nearest chinese restaurant but also a traffic jam that needs to be avoided between here and there. This is a composition problem.

Intelligence applied to the viewer input consists of active display changing whether it be the state of Super Mario, the straightening of hand drawn lines, or text completion in a phone book look up. Information utility is what is attractive on interactive TV. If the utility is only provided by viewer input simplification, utility would be low. This is like a pong game without a bounce. The data source dominates the utility. How the automaton is intelligent about the data source is most critical. Like any idea that can now be discussed, the idea of data automatons is old. "Intelligent agents" were deeply analysed in the early 1970s. Making these computerized agents work with high utility values is a different matter.

Agents are best that accomplish the narrowest function. The best appear to be incredibly simple services that give incredibly good information. An example would be the traffic-to-work service. The set-up is to indicate that you want this to work all the time for as long as you do not trash it. Then you indicate your location and the office location (perhaps by street addresses). This is the I_{self} cost (essentially 3 bits). The automaton multiplier, **A**, is enormous since the automaton does the customization, the I_{other} value is high because somebody wrote a fairly complicated computer program, and the "think cost," **T**, is essentially fractional. The result is an extremely high utility value for any information event which now occurs with this information service.

In effect, the conclusion is that we should not use interactive TV to get greedy and have people shop by the instant. People don't like to make a lot of decisions. Rather we should use interactive TV to create the ultimate narrowcasts. This can be done with good content producers and with a broadcast structure underpinning interactive TV. Within the communities of the two key types of content producers, entertainment artists and software automaton producers, there are incredibly gifted people who want to strut their stuff. But don't forget that nearly everybody knows something of great interest to a respectable number of other people. Smart TV shouldn't leave anyone behind.

TO PROBE FURTHER. Some of the earliest descriptions of the practical requirements for home information services published after the arrival (and disappointment) of Videotex can be found in *Computer Message Systems* (McGraw-Hill, New York, 1984), Jacques Vallee and *Viewdata and the Information Society* (Prentice-Hall, Englewood Cliffs, NJ, 1982), James Martin. A forward looking book on the potential of smart TV is George Gilder's *Life After Television* (Norton, New York, London), 1992. For an advertiser's perspective on interactive television, see *Cable TV Advertising* (Quorum Books, Westport, CT, 1989), Rajeev Batra and Rashi Glazer. W.

Wayne Talarzyk and Robert E. Widing, III have much to say about what consumers find to be valuable and how technology can enhance the value of information in Chapter 9 of *Telecommunications, values, and the public interest* (Ablex Publishing, Norwood, NJ, 1990), editor, Sven B. Lundsted. Chapter 10 of the same text, by Marilyn Greenwald provides some suggestions as to why videotex hasn't yet reached its potential. Interesting announcements related to smart TV appear almost daily in major newspapers. For a recent survey of the alliances which have been formed to advance this technology is *Vague New World: Digital Media Business Takes Form as a Battle Of Complex Alliances*, The Wall Street Journal, July 14, 1993, A1:6, by Stephen Kreider Yoder and G. Pascal Zachary, or *Battles loom for control of TV's portal to cable*, New York Times, April 3, 1993, A, 43:5, by John Markoff. The Free Software Foundation is described in: *Programs to the People*, Technology Review, v94n2, 52-60, Feb 1991, Simson L. Garfinkel. Gophers and the related Wide Area Information Servers (WAIS) are summarized by David Churbuck in *Good-bye, Dewey decimals* (Forbes, v151n4, 204-205, February 15, 1993). A related paper on information cost structure in user interfaces is Russell, D. M., Stefik, M. J., Pirolli, P., and Card, S. K. *The cost structure of sensemaking.* Proceedings of INTERCHI, 1993 (Amsterdam, The Netherlands, 24-29 April 1993), pp. 269-276. Also see, Card, S. K., Robertson, G. G., & Mackinlay, J. D. *The information visualizer, an information workspace.* Proceedings of the ACM Conference on Human Factors in Software. New York: ACM, pp. 181-188, 1991.

Biographical Annotation

Robert Thibadeau has been Director of the Robotics Institute's Imaging Systems Laboratory in the School of Computer Science at Carnegie Mellon University since 1981. With a diverse background in computational linguistics, experimental cognitive science, applied computer vision, artificial intelligence languages, computer aided design and manufacturing, document image management and conversion, chip and electronics design, he has authored dozens of systems, including many industrial and commercial systems, around technologies of highly skilled perception. The laboratory is one of the founding laboratories of the Robotics Institute and is currently one of the laboratories working for the Advanced Research Projects Agency on High Definition Display Technology (ARPA order 6873, contract #MDA972-92-J-1010). This is related work on the topic of "intelligent frame buffers." Dr. Thibadeau is author of a book "The Television Computer" available from Visual Understanding Systems, Inc., (412) 488-3600, or by fax - 3611, for $39.95 (+$4.50 S/H). Dr. Thibadeau's group also has patents pending in several areas including certain location-based information filters as improvements to TV broadcast networks.